Zoosyst. Evol. 101 (1) 2025, 17–34 | DOI 10.3897/zse.101.136882

A new species of *Semisulcospira* O. Boettger, 1886 (Gastropoda, Cerithoidea, Semisulcospiridae) from Fujian, China with mitochondrial genome and its phylogenetic implications

Yi-Bin Xu[1*], Yuan-Zheng Meng[2*], Sheng Zeng[2*], Hang-Jun Wang[3,4], Shen Zhong[5], De-Yuan Yang[2,6], Xi-Ping Zhou[7], Christopher J. Glasby[8,9]

1 *Key Laboratory of Cultivation and High-value Utilization of Marine Organisms in Fujian Province, Fisheries Research Institute of Fujian, Xiamen 361102, China*
2 *College of the Environment and Ecology, Xiamen University, Xiamen 361102, China*
3 *Wenzhou Marine Center, Ministry of Natural Resources of the People's Republic of China, Wenzhou, China*
4 *Marine Ecosystem Observation and Research Station on the Yangtze River Estuary, Wenzhou 325011, China*
5 *Xiamen Ocean Vocational College, Xiamen 361102, China*
6 *National Taiwan Ocean University, Keelung, Taiwan, China*
7 *Department of Environmental Science and Engineering, Xiamen University Tan Kah Kee College, Zhangzhou 363105, China*
8 *Natural Sciences, Museum & Art Gallery Northern Territory, PO Box 4646, Darwin, NT 0801, Australia*
9 *Australian Museum Research Institute, 1 William Street, NSW 2010, Sydney, Australia*

https://zoobank.org/F686B6D9-0DE3-4BB4-94E5-26FEA80E64CD

Corresponding authors: De-Yuan Yang (deyuanyang92@163.com); Xi-Ping Zhou (zoeyzhou@xujc.com); Christopher J. Glasby (glasby93@gmail.com)

Academic editor: Frank Köhler ♦ **Received** 11 September 2024 ♦ **Accepted** 10 November 2024 ♦ **Published** 2 January 2025

Abstract

During our investigation of the diversity and phylogeny of the family Semisulcospiridae in China, we have found a presumed new species in mountain streams of Xiamen, Fujian Province on both morphology and molecular evidence. This new species, *Semisulcospira egretta* **sp. nov.**, is characterized by its thin and smooth shell, retaining about 0.5 mm long embryonic shells in its brood pouch, and by distinctive features of its radular morphology. Phylogenetic analyses of partial sequences of the mitochondrial markers 16S ribosomal RNA (*16S*) and cytochrome c oxidase subunit I (*COI*) and complete sequences of the nuclear ribosomal internal transcribed spacer (ITS, including internal transcribed spacer 1, 5.8S ribosomal RNA gene, and internal transcribed spacer 2) supported the distinctiveness of *S. egretta* **sp. nov.** from other *Semisulcospira* species. These analyses consistently placed *S. egretta* **sp. nov.** in *Semisulcospira* confirming its monophyly. Phylogenetic analyses based on just *COI* and *16S* exhibited low resolution rendering the phylogenetic position of the new species within the genus uncertain. In contrast, phylogenetic tree analyses of complete mitochondrial genomes, which include 13 protein-coding genes (PCGs), were better resolved. However, the mitogenome dataset included a more restricted sampling of taxa (16), impeding a detailed exploration of the phylogenetic position of the new species. To address this deficit, we provide seven complete mitochondrial genomes for *S. egretta* **sp. nov.**, obtained using the genome skimming sequencing method. This study contributes important molecular and genomic data, aiding in the understanding of semisulcospirid phylogeny, and highlights the need for further taxonomic sampling and phylogenomic methods to resolve the evolutionary relationships within Semisulcospiridae.

Key Words

16S rRNA, *COI*, mitogenome, phylogenetic analysis, semisulcospirid gastropods

* These authors contributed equally to this work.

Copyright *Xu, Y.-B. et al.* This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Introduction

Semisulcospiridae Morrison, 1952 is a family of freshwater gastropods comprising more than 90 species from four genera (Lydeard and Cummings 2019; Strong et al. 2022; Du and Yang 2023; MolluscaBase eds. 2024). This family is predominantly found in East Asia and North America, with the majority (over 60 species from three genera) documented in China (Du et al. 2019a, 2019b; Du and Yang 2023; He et al. 2024).

Semisulcospira O. Boettger is the type genus of the family, characterized by its viviparous reproductive strategy, while the other three genera are oviparous (Köhler 2017; Du et al. 2019b; Strong et al. 2022). Members of *Semisulcospira* are mainly distributed in East Asia and prefer living in freshwater streams, rivers, and lakes (Lydeard and Cummings 2019; Strong et al. 2022; Du and Yang 2023). They are important to environmental studies and public health (Köhler 2016); however, our understanding of their taxonomy and phylogenetic systematics remains limited. Due to significant morphological variation in shell characteristics, *Semisulcospira* contains numerous nominal species. Recently, anatomical and molecular studies have demonstrated that many species of Semisulcospiridae are misidentifications, even at the genus level (Du et al. 2019a, 2019b; Strong et al. 2022; Sawada et al. 2024). For example, the genera *Namrutua* Abbott, 1948 and *Senckenbergia* Yen, 1939 have been proven synonyms of *Semisulcospira*, and *Melania dulcis* Fulton, 1904, *M. lauta* Fulton, 1904, and *Semisulcospira inflata* S. Tchang, 1949 have been regarded as synonyms of *Hua textrix* (Heude, 1889) (Du et al. 2019a; Du et al. 2019b).

Currently, the genus *Semisulcospira* comprises more than 36 valid species, with 18 endemic to Lake Biwa in Japan. Approximately 10 species have been documented from China, with *Semisulcospira libertina* (A. Gould, 1859) as the only species recorded in Fujian Province (Liu et al. 1979; Liu et al. 1994; Du et al. 2019a, 2019b; Du and Yang 2023; MolluscaBase eds. 2024; Sawada et al. 2024). As molecular methods have been introduced in the systematic study of semisulcospirids, numerous cryptic species have been discovered in China, Japan, and the U. S. (Du et al. 2019a; Sawada and Nakano 2021; Strong et al. 2022; Chen et al. 2023; Du and Yang 2023; Sawada et al. 2024). Previous research has demonstrated high diversity within this family, indicating the potential for more cryptic species (Strong et al. 2022; Du and Yang 2023; He et al. 2024).

During the investigation of the diversity and phylogeny of Semisulcospiridae in China, we found a presumed new species from mountain streams of Xiamen, Fujian Province, based on both morphological and molecular data. Additionally, we provided seven complete mitochondrial genome (mitogenome) sequences of the putative new species using the genome skimming sequencing method (Yang et al. 2024). Phylogenetic analysis was performed using two partial mitochondrial genes, cytochrome oxidase subunit 1 (*COI*) and 16S ribosomal RNA (*16S*), as well as the complete mitogenomes.

Material and methods

Specimens and sampling

Samples were collected from streams in Xiamen, Fujian Province, China from 2022 to 2023. Examined specimens comprised dry shells and ethanol-preserved samples deposited in the Marine Biological Sample Museum, Third Institute of Oceanography, Ministry of Natural Resources (Xiamen, China) and De-Yuan Yang's laboratory, College of the Environment and Ecology, Xiamen University, Xiamen, Fujian Province, China (Fig. 1).

Morphological methods

Living specimens for dissecting were soaked in hot water for 1 minute to separate soft parts from the shells. Anatomical techniques and terminology follow Strong (2011) and Strong and Köhler (2009). Specimens were photographed with an Olympus E-M1 Mark II camera with a 60 mm macro lens, and image stacks were obtained using Helicon Focus 7 (https://www.heliconsoft.com/heliconsoft-products/helicon-focus/) and post-processed using Adobe Photoshop CC 2019. Shells were photographed from two different perspectives, apertural view (viewed from aperture with shell parallel to the coiling axis) and abapertural view (viewed opposite from the apertural view). The morphological parameters of the shells were measured using calipers with a precision of 0.1 mm, following Du et al. (2019a). The radulae and embryos were digested by Proteinase K at 55 °C for 2 hours, and then cleaned, dried, coated with gold, and imaged under SEM (Phenom Prox). For more information about the specimens used in this study (Suppl. material 1: table S1).

Abbreviations

MBSM: the Marine Biological Sample Museum, Third Institute of Oceanography, Ministry of Natural Resources (Xiamen, China). **XMU_DYY:** Laboratory of De-Yuan Yang, College of the Environment and Ecology, Xiamen University (Xiamen, Fujian Province, China). **ZMB:** Museum für Naturkunde, Berlin. **NHMUK:** Natural History Museum, London. **MNHN:** Muséum national d'Histoire naturelle, Paris. **YZM:** Collection by Yuan-Zheng Meng. **LWL:** Collection by Li-Wen Lin. **SEM:** Scanning electron microscope. ***16S*:** 16S ribosomal RNA. ***COI*:** Cytochrome oxidase subunit I. **ITS:** Internal transcribed spacer complete sequence, including internal transcribed spacer 1, 5.8S ribosomal RNA gene, and internal transcribed spacer 2. **PCGs:** Protein-coding genes. **ML:** Maximum likelihood. **BI:** Bayesian inference. **PP:** Posterior probability. **BS:** Ultrafast bootstrap support. **H:** Shell height. **B:** Shell width. **BW:** Height of the last whorl. **WA:** Width of aperture. **LA:** Length of aperture.

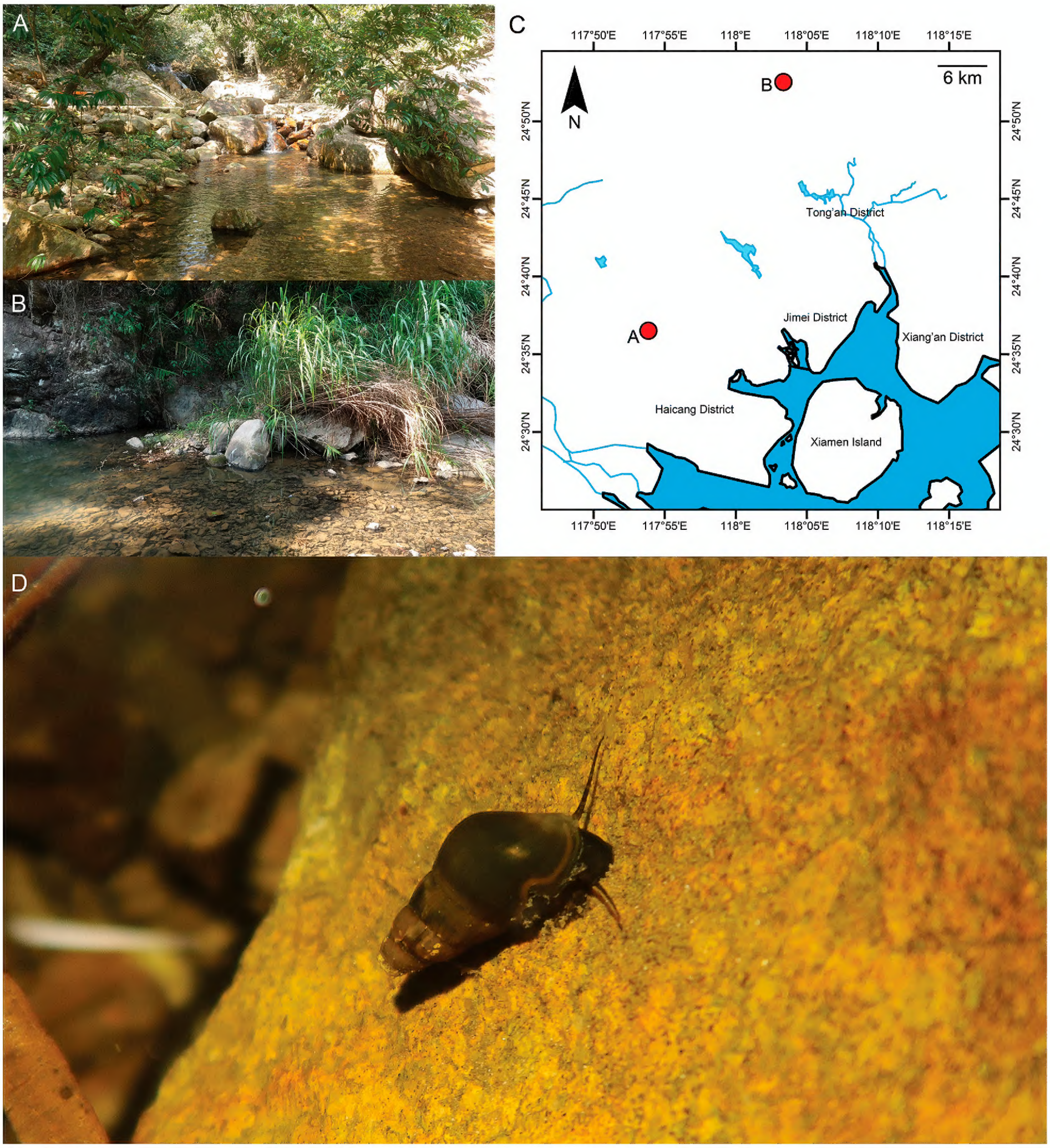

Figure 1. The habitat and sampling localities of *S. egretta* sp. nov. in Xiamen City, Fujian Province, China. **A.** Type locality, 24°36'29"N, 117°53'39"E, in Haicang District; **B.** 24°52'35"N, 118°03'01"E, in Tong' an District; **C.** A simplified map of Xiamen, locality A and B are marked with red points; **D.** A living specimen of *S. egretta* sp. nov. All photos are taken by Yuan-Zheng Meng.

DNA extraction, amplification, and sequencing

Genomic DNA was extracted from a 1 mm^3 sample of muscle tissue from the foot of specimens using a TIANamp Genomic DNA Kit (TIANGEN, Beijing, China) according to the manufacturer protocols.

Two partial mitochondrial markers:16S rRNA (*16S*, 467 bp) and cytochrome oxidase subunit I (*COI*, 655 bp) were sequenced from three specimens of the new species (XMU_DYY_XM02, XM04, and XM05). For detailed information on protocols and primers of PCR (polymerase chain reaction), see Suppl. material 1: table S2.

Mitogenome sequences were obtained from seven specimens (XMU_DYY_XMA03, XMA05, XMA06, XMB04, XMC01, XMU_DYY_RTM05, and RTM25). Genome skimming for the mitogenome sequences was conducted using next-generation sequencing (NGS) on the Illumina NovaSeq X Plus and DNBSEQ-T7 platforms, generating paired-end reads of 150 bp in length. The sequencing was performed by Novogene Bioinformatics Technology Co., Ltd. in Beijing, China.

Mitogenome assembly and annotation

The assembly and annotation strategy of mitogenomes followed the methods of Yang et al. (2024) and Xu et al. (2024). In summary, we initially assembled the data using Getorganelle v1.7.6.1 (Jin et al. 2020). If assembly failed, we used other Semisulcospiridae mitogenomes as seeds for reassembly.

Nuclear genes assembly and annotation

To evaluate the *p*-distance of internal transcribed spacer (ITS) sequences, we extracted these sequences from genome skimming data of the seven newly sequenced specimens, following the protocol outlined by Yang et al. (2024). Briefly, we employed Getorganelle v1.7.6.1 (Jin et al. 2020), using all publicly available 18S ribosomal RNA (18S rRNA), 28S ribosomal RNA (28S rRNA), and ITS sequences from the family Semisulcospiridae as seed sequences. In addition, we assembled all publicly available genome skimming data from the NCBI Sequence Read Archive (SRA accession numbers: SRR26074378, SRR26107127, SRR12518401) to obtain ITS gene sequences, and successful assembly was achieved for the data SRR26107127 and SRR12518401 (Suppl. material 1: table S3).

Molecular analyses

The phylogenetic analysis was conducted in four parts: 1) three separate datasets—a concatenated *16S*+*COI* dataset integrated from previous studies containing 132 individuals, each with both *16S* and *COI* sequences, as well as separate *16S* and *COI* datasets (Suppl. material 1: table S3); 2) the extensive *COI* dataset (*COI*-ex), comprising 877 sequences (Suppl. material 1: table S4); 3) the ITS dataset, including 9 newly obtained ITS sequences in this study and two existing ITS sequences in NCBI (Suppl. material 1: table S4); 4) the mitogenome dataset, including the seven newly sequenced mitogenomes of *S. egretta* sp. nov. and the dataset from Xu et al. (2024) (Suppl. material 1: table S5). *Hua jacqueti* (Dautzenberg & H. Fischer, 1906) and *Juga plicifera* (I. Lea, 1838) were used as the outgroups of datasets 1 & 2, following Strong and Köhler (2009) and Köhler (2016). Sequences used in this study were integrated from previously published studies (Strong and Köhler 2009; Zeng et al. 2015; Hilgers et al. 2016; Köhler 2016; Whelan and Strong 2016; Zeng et al. 2016; Chiu et al. 2017; Köhler 2017; Kim and Lee 2018; Nguyen et al. 2018; Du et al. 2019a, 2019b; Guo et al. 2019; Hartnell College Genomics Group et al. 2019; Lee et al. 2019; Stelbrink et al. 2019; Xu et al. 2019; Lee et al. 2020; Miura et al. 2020; Yan et al. 2020a, b; Choi et al. 2021; Gim et al. 2021; Kato et al. 2022; Ling et al. 2022; Strong et al. 2022; Yang and Deng 2022; Yin et al. 2022; He et al. 2024; Xu et al. 2024).

Most subsequent analyses were conducted in PhyloSuite v1.2.3 (Zhang et al. 2020), including multiple sequence alignment, trimming, model selection, and phylogenetic tree construction.

For the *16S* and *COI* sequences, we initially extracted the full-length *16S* and *COI* sequences from our newly assembled mitochondrial genomes and used them as reference sequences for alignments. The alignments were performed using the online version of MAFFT (https://mafft.cbrc.jp/alignment/server/, accessed October 2024) with default settings. The following parameters were adjusted: UPPERCASE/lowercase: same as input; Direction of nucleotide sequences: adjusted according to the first sequence (sufficient for most cases); Output order: same as input.

For the mitogenomes dataset, we only focused on the 13 protein-coding genes (PCGs). These sequences were aligned using the MAFFT v.7.313 (Katoh and Standley 2013) plugin in PhyloSuite under the codon alignment mode. All aligned sequences were inspected in Geneious Prime 2024.0.4 (https://www.geneious.com) before and after trimming.

Considering the *COI* and *16S* sequences mainly from Köhler (2016, 2017) and Du et al. (2019a, 2019b), which were amplified using different primers, leading to differences in the amplified regions (Fig. 2), we applied five trimming methods for dataset 1 to to assess the impact of trimming on phylogenetic reconstruction: (1) using TrimAl v.1.2 (Capella-Gutiérrez et al. 2009) with the automated1 setting (referred to as auto); (2) trimming sequences based on the length of sequences from Du et al. (2019b, 2019b) as a reference (referred to as 'du'); (3) trimming sequences based on the length of sequences from Köhler (2016, 2017) as a reference (referred to as 'ko'); (4) using the shared overlapping region between Köhler (2016, 2017) and Du et al. (2019b, 2019b) as a reference ('du ∩ ko'); and (5) using the combined region from both Köhler (2016, 2017) and Du et al. (2019b, 2019b) as a reference ('du ∪ ko') (Fig. 2). As a result, the dataset 1 generated 15 sub-datasets (see Suppl. materials 2–4). Datasets 2–4 were trimmed under the auto mode (Suppl. material 1: table S6).

ModelFinder v.2.2.0 (Kalyaanamoorthy et al. 2017) was used to select the best substitution models under the Bayesian Information Criterion (BIC) for maximum likelihood (ML) analysis and the corrected Akaike Information Criterion (AICc) for Bayesian inference (BI) analysis (Suppl. material 1: table S7). Maximum likelihood phylogenetic reconstruction was conducted in IQ-TREE v.2.2.2 (Nguyen et al. 2015), using the best partition scheme and under an edge-linked partition model for 20000 ultrafast bootstraps (UFBoot2). Bayesian Inference phylogenies were inferred using MrBayes v.3.2.7a (Ronquist et al. 2012) under a partition model (2 parallel runs, 2 000 000 generations), ensuring the average standard deviation of split frequencies (ASDSF) value was below 0.01 (Ronquist et al. 2012). If convergence was not achieved after 2,000,000 generations (ASDSF value > 0.01), additional generations were added to continue the analysis.

Figure 2. A schematic diagram of the *COI* gene trimming methods.

Pairwise tree structure comparison was conducted using the all.equal.phylo function in the ape v.5.7.1 package (Paradis and Schliep 2019), and topological differences were assessed using TreeSpace (Jombart et al. 2017), both implemented in R v.4.3.1 (R Core Team 2023). Finally, iTOL v.6 (Letunic and Bork 2024) was used to visualize the trees.

Summary statistics for multiple alignments were generated using a custom Python script, which utilizes the alignment summary function in BioKIT (Steenwyk et al. 2022) (Table 1).

Genetic distances (*p*-distance and Kimura 2 Parameters, K2P) were computed using MEGA X with default settings (Kumar et al. 2018). A heatmap of the *p*-distances between selected specimens (clade D of Fig. 3) was generated by TBtools (Chen et al. 2020) (Fig. 4, Suppl. material 1: table S8).

Table 1. Summary statistics for multiple alignment of various datasets.

Dataset	Gene	Alignment length	Trimal methods	Trimal length	Percentage (%)	Constant sites	Number of taxa	Parsimony informative sites	Variable sites
Mitogenome	*ATP6*	705	auto	693	98.3	267	32	355	426
	ATP8	177	auto	159	89.8	39	32	113	120
	COX1	1533	auto	1512	98.6	882	32	563	630
	COX2	687	auto	666	96.9	327	32	281	339
	COX3	777	auto	759	97.7	354	31	335	405
	CYTB	1137	auto	1134	99.7	526	31	539	608
	ND1	942	auto	933	99	390	32	467	543
	ND2	1077	auto	1062	98.6	298	32	663	764
	ND3	351	auto	324	92.3	130	32	172	194
	ND4	1380	auto	1359	98.5	442	32	810	917
	ND4L	321	auto	288	89.7	101	31	144	187
	ND5	1719	auto	1689	98.3	599	32	933	1090
	ND6	549	auto	492	89.6	136	32	307	356
	Concatenation	11355	auto	11070	97.5	4470	32	5702	6600
COI+16S	*COI*	1538	auto	1533	99.7	1018	132	469	515
			du	724	47.1	429	132	277	295
			ko	658	42.8	368	132	272	290
			du ∩ ko	550	35.8	312	132	225	238
			du ∪ ko	832	54.1	485	132	324	347
	16S	1459	auto	1345	92.2	791	132	460	554
			du	544	37.3	333	132	176	211
			ko	817	56.0	410	132	348	407
			du ∩ ko	527	36.1	324	132	171	203
			du ∪ ko	834	57.2	419	132	353	415
	COI+16S	2997	auto	2878	96.0	1809	132	929	1069
			du	1268	42.3	762	132	453	506
			ko	1475	49.2	778	132	620	697
			du ∩ ko	1077	35.9	636	132	396	441
			du ∪ ko	1666	55.6	904	132	677	762
COI-ex	*COI*	1533	auto	548	35.7	279	8877	247	270
ITS	ITS	2657	auto	964	36.3	804	11	69	160

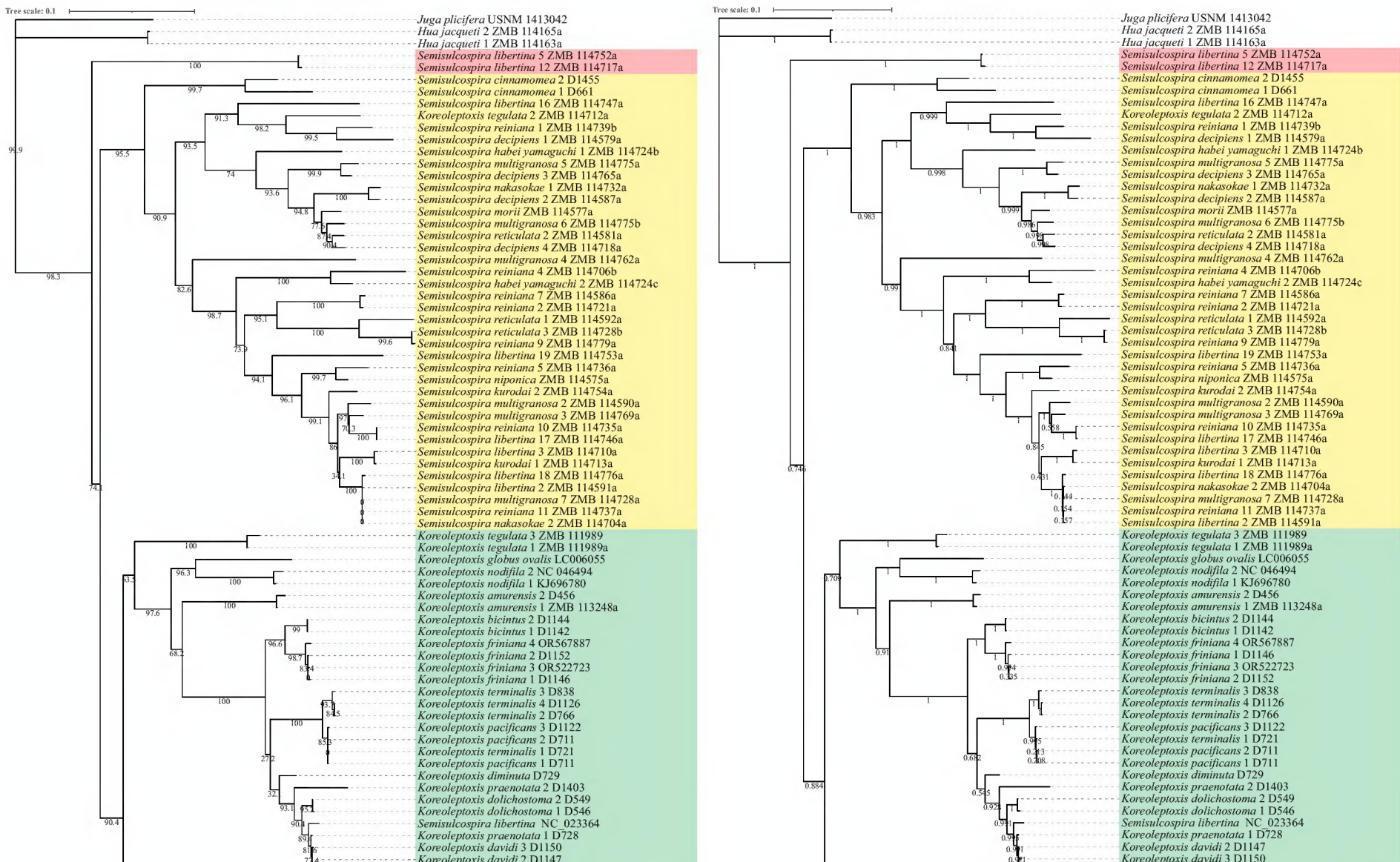
Tree scale: 0.1
Juga plicifera USNM 1413042
Hua jacqueti 2 ZMB 114165a
Hua jacqueti 1 ZMB 114163a
Semisulcospira libertina 5 ZMB 114752a
Semisulcospira libertina 12 ZMB 114717a
Semisulcospira cinnamomea 2 D1455
Semisulcospira cinnamomea 1 D661
Semisulcospira libertina 16 ZMB 114747a
Koreoleptoxis tegulata 2 ZMB 114712a
Semisulcospira reiniana 1 ZMB 114739b
Semisulcospira decipiens 1 ZMB 114579a
Semisulcospira habei yamaguchi 1 ZMB 114724b
Semisulcospira multigranosa 5 ZMB 114775a
Semisulcospira decipiens 3 ZMB 114765a
Semisulcospira nakasokae 1 ZMB 114732a
Semisulcospira decipiens 2 ZMB 114587a
Semisulcospira morii ZMB 114577a
Semisulcospira multigranosa 6 ZMB 114775b
Semisulcospira reticulata 2 ZMB 114581a
Semisulcospira decipiens 4 ZMB 114718a
Semisulcospira multigranosa 4 ZMB 114762a
Semisulcospira reiniana 4 ZMB 114706b
Semisulcospira habei yamaguchi 2 ZMB 114724c
Semisulcospira reiniana 7 ZMB 114586a
Semisulcospira reiniana 2 ZMB 114721a
Semisulcospira reticulata 1 ZMB 114592a
Semisulcospira reticulata 3 ZMB 114728b
Semisulcospira reiniana 9 ZMB 114779a
Semisulcospira libertina 19 ZMB 114753a
Semisulcospira reiniana 5 ZMB 114736a
Semisulcospira niponica ZMB 114575a
Semisulcospira kurodai 2 ZMB 114754a
Semisulcospira multigranosa 2 ZMB 114590a
Semisulcospira multigranosa 3 ZMB 114769a
Semisulcospira reiniana 10 ZMB 114735a
Semisulcospira libertina 17 ZMB 114746a
Semisulcospira libertina 3 ZMB 114710a
Semisulcospira kurodai 1 ZMB 114713a
Semisulcospira libertina 18 ZMB 114776a
Semisulcospira libertina 2 ZMB 114591a
Semisulcospira multigranosa 7 ZMB 114728a
Semisulcospira reiniana 11 ZMB 114737a
Semisulcospira nakasokae 2 ZMB 114704a
Koreoleptoxis tegulata 3 ZMB 111989
Koreoleptoxis tegulata 1 ZMB 111989a
Koreoleptoxis globus ovalis LC006055
Koreoleptoxis nodifila 2 NC 046494
Koreoleptoxis nodifila 1 KJ696780
Koreoleptoxis amurensis 2 D456
Koreoleptoxis amurensis 1 ZMB 113248a
Koreoleptoxis bicintus 2 D1144
Koreoleptoxis bicintus 1 D1142
Koreoleptoxis friniana 4 OR567887
Koreoleptoxis friniana 2 D1152
Koreoleptoxis friniana 3 OR522723
Koreoleptoxis friniana 1 D1146
Koreoleptoxis terminalis 3 D838
Koreoleptoxis terminalis 4 D1126
Koreoleptoxis terminalis 2 D766
Koreoleptoxis pacificans 3 D1122
Koreoleptoxis pacificans 2 D711
Koreoleptoxis terminalis 1 D721
Koreoleptoxis pacificans 1 D711
Koreoleptoxis diminuta D729
Koreoleptoxis praenotata 2 D1403
Koreoleptoxis dolichostoma 2 D549
Koreoleptoxis dolichostoma 1 D546
Semisulcospira libertina NC_023364
Koreoleptoxis praenotata 1 D728
Koreoleptoxis davidi 3 D1150
Koreoleptoxis davidi 2 D1147
Koreoleptoxis davidi 1 D1143
Semisulcospira pleuroceroides 2 D1043
Semisulcospira pleuroceroides 1 D1042
Semisulcospira ningpoensis 3 D1022
Semisulcospira ningpoensis 2 D656
Semisulcospira ningpoensis 1 D1077
Semisulcospira gredleri 4 D1272
Semisulcospira gredleri 2 D663
Semisulcospira gredleri 1 D180
Semisulcospira gredleri 5 D1453
Semisulcospira gredleri 3 D1052
Semisulcospira egretta sp. nov. 18 XMB04
Semisulcospira egretta sp. nov. 2 RTM25
Semisulcospira egretta sp. nov. 19 XMC01
Semisulcospira egretta sp. nov. 11 TZS_008
Semisulcospira egretta sp. nov. 6 TZS_002
Semisulcospira egretta sp. nov. 7 MBSM-2024-XM-001
Semisulcospira egretta sp. nov. 10 TZS_007
Semisulcospira egretta sp. nov. 16 XMA05
Semisulcospira egretta sp. nov. 17 XMA06
Semisulcospira egretta sp. nov. 1 RTM05
Semisulcospira egretta sp. nov. 15 XMA03
Semisulcospira egretta sp. nov. 13 TZS_010
Semisulcospira egretta sp. nov. 9 TZS_006
Semisulcospira egretta sp. nov. 5 XM02
Semisulcospira egretta sp. nov. 4 XM05
Semisulcospira egretta sp. nov. 8 TZS_005
Semisulcospira egretta sp. nov. 14 TZS_011
Semisulcospira egretta sp. nov. 3 XM04
Semisulcospira egretta sp. nov. 12 TZS_009
Semisulcospira dolorosa ZMB 102182a
Semisulcospira libertina 6 ZMB 114751a
Semisulcospira libertina 15 ZMB 114716a
Semisulcospira reiniana 6 ZMB 114583a
Semisulcospira libertina 1 ZMB 114589a
Semisulcospira libertina 8 ZMB 114740a
Semisulcospira libertina 14 ZMB 114703b
Semisulcospira libertina 13 ZMB 114703a
Semisulcospira multigranosa 1 ZMB 114584a
Semisulcospira libertina 7 ZMB 114739a
Semisulcospira libertina 4 ZMB 114711a
Semisulcospira libertina 11 ZMB 114705a
Semisulcospira libertina 10 ZMB 114780a
Semisulcospira reiniana 3 ZMB 114738a
Semisulcospira reiniana 8 ZMB 114708a
Semisulcospira libertina 9 ZMB 114734a
Semisulcospira calculus 2 D1592
Semisulcospira calculus 1 D1594
Semisulcospira coreana 4 NC 037771
Semisulcospira coreana 2 ZMB 114761a
Semisulcospira libertina 20 ZMB 114758a
Semisulcospira coreana 1 ZMB 114720a
Semisulcospira forticosta 2 ZMB 114715a
Semisulcospira forticosta 1 ZMB 111991a
Semisulcospira gottschei 5 MK559478
Semisulcospira gottschei 6 ZMB 112961a
Semisulcospira gottschei 1 ZMB 111992a
Semisulcospira gottschei 4 ZMB 114781
Semisulcospira coreana 3 ZMB 114731a
Semisulcospira gottschei 3 ZMB 114709a
Semisulcospira gottschei 2 ZMB 114745a
Clade A
Clade B
Clade C
Clade D

Figure 3. Maximum likelihood tree (left) and Bayesian inference tree (right) inferred from *COI* and *16S* gene sequences. Bootstrap supports (BS) /posterior probabilities (PP) are shown on the nodes. Notes: *Semisulcospira egretta* sp. nov. is highlighted in red.

Results

DNA barcoding

The average interspecific genetic distances (*p*-distance) for *COI* and *16S* sequences between *Semisulcospira egretta* sp. nov. and other congeneric species ranged from 11.9% to 14.4%, and 5.0% to 7.4%, respectively. This value is significantly higher than the average intraspecific genetic distances within *S. egretta* sp. nov., which varied from 0% to 3.1% (0.84% on average for *COI*) and 0% to 2.2% (0.73% on average for *16S*). The *p*-distance of ITS sequences between *Semisulcospira egretta* sp. nov ranged from 0% to 1% (0.56% on average) (Fig. 4, Suppl. material 1: table S8).

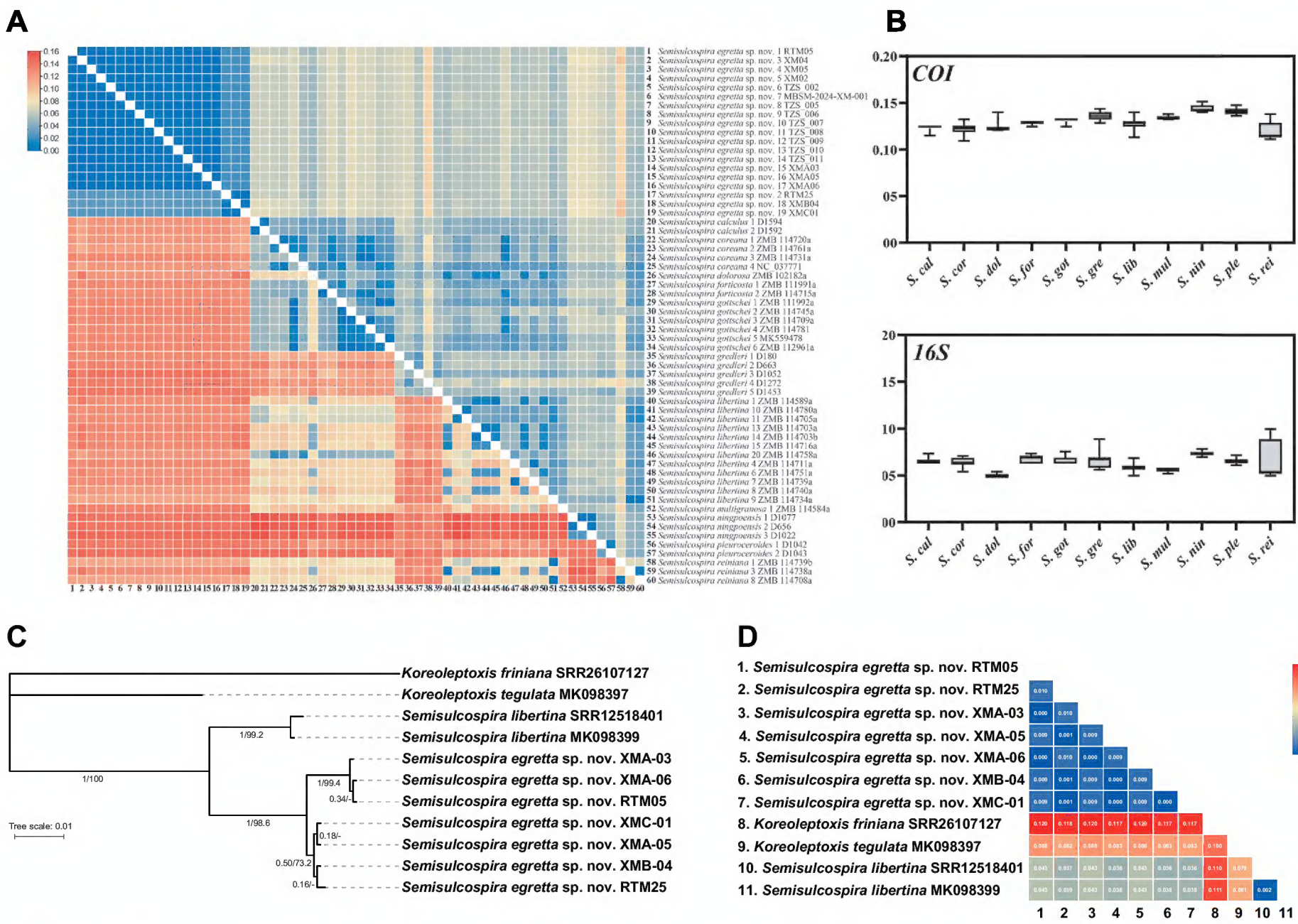

Figure 4. A. A heatmap of *p*-distances between selected congeners. The *p*-distances based on *COI* and *16S* are shown at the bottom left and upper right, respectively; **B.** Box plots of interspecific *p*-distances between *S. egretta* sp. nov. and other species based on *COI* (upper) and *16S* (bottom). The specific name of the congeners is indicated by the first three letters; **C.** ML and BI phylogenetic trees of ITS genes; **D.** *p*-distance of *S. egretta* sp. nov. and other semisulcospirids.

Phylogenetic analyses

A summary of the characteristics of the data matrices can be found in Table 1. The mitogenome dataset (13 PCGs) includes 16 Semisulcospiridae sequences, with a total length of 11070 bp and 5702 parsimony informative sites. In the *COI+16S* dataset, we selected the sub-dataset (du ∪ ko) for further analysis. This sub-dataset includes sequences from 132 individuals, with a total length of 1666 bp and 677 parsimony informative sites. The trimmed *COI* dataset (du ∪ ko) is 832 bp in length, containing 324 parsimony informative sites, and the *16S* dataset (du ∪ ko) is 834 bp in length, containing 353 parsimony informative sites. The trimmed *COI*-ex dataset consists of sequences from 863 taxa that are 548 bp in length and include 247 parsimony informative sites.

Phylogenetic reconstruction of the *16S+COI* dataset, using different trimming methods and employing Maximum Likelihood (ML) and Bayesian Inference (BI), resulted in 30 phylogenetic trees. TreeSpace (Jombart et al. 2017) was used to assess potential discrepancies among these trees. Analyzing all 30 trees in TreeSpace showed that the concatenated *16S+COI* dataset exhibited fewer discrepancies compared to the individual *16S* or *COI* datasets (Suppl. materials 2–4). For the *16S* dataset, inferred trees based on different trimming methods did not cluster together (Suppl. material 5). In the *COI* dataset, three tree clusters were identified. Phylogenetic trees inferred using ML from different trimmed sub-datasets formed one cluster, whereas trees inferred using BI were divided into two distinct clusters (Suppl. material 5).

In the combined *16S+COI* dataset, three tree clusters were identified. Generally, phylogenetic trees derived from different trimming methods using the same inference approach clustered together, with the exception of the ML tree from the 'du' sub-dataset (Suppl. material 5).

Based on this evidence, we can conclude that the phylogenetic tree inferred from the combined *16S+COI* dataset is more reliable compared to those inferred from the individual *16S* or *COI* datasets. Additionally, we propose that different trimming methods have an influence on the phylogenetic reconstruction.

Given that the ML and BI trees of the sub-dataset (du ∪ ko) cluster together closely, we present these two trees in more detail.

Phylogenetic trees (*16S+COI*) using maximum likelihood (ML) and Bayesian Inference (BI) are both divided

into four primary clades. Clades A, B, and D only include *Semisulcospira* species, while clade C only includes *Koreoleptoxis* species (except NC_023364, see below). *S. egretta* sp. nov. forms a distinct lineage inside clade D (PP = 0.547, BS = 92.4%).

Phylogenetic analyses (ML and BI) inferred from the *COI* datasets have different topologies (Suppl. material 4), but *Semisulcospira egretta* sp. nov. has a close relationship with a clade containing 31 sequences of eight species (*S. gottschei*, *S. coreana*, *S. forticosta*, *S. libertina*, *S. calculus*, *S. multigranosa*, *S. reiniana*, and *S. dorlosa*, PP = 0.58; BS = 92.8%). Meanwhile, phylogenetic trees inferred from the *16S* dataset show that *S. egretta* sp. nov. have a close relationship with *Semisulcospira multigranosa* (O. Boettger, 1886) in the ML and BI tree (PP = 0%, BI = 0.36, Suppl. material 4). Phylogenetic analyses inferred from the *COI*-ex (ML) dataset and the ITS dataset (ML and BI) all showed that *S. egretta* formed a monophyletic clade (BS = 100% in *COI*-ex, PP = 1, BS = 98.6% in ITS, Suppl. material 6 and Fig. 4C).

The topologies of the phylogenetic trees inferred from 13 PCGs are broadly consistent with the phylogenetic tree in Xu et al. (2024) (Fig. 5A). The tree contains 30 Cerithoidea mitogenome sequences, 16 of which are Semisulcospiridae sequences: 10 *Semisulcospira* (including the 7 newly reported ones in this study), five *Koreoleptoxis*, and one *Hua*. Nine *Semisulcospira* sequences clustered in a clade, including the 7 newly reported sequences of *S. egretta* sp. nov. The new species is closely related to *S. gottschei* and *S. coreana*, with high support values (BS = 100%, PP = 1). Another clade, a sister group of the former, contains five *Koreoleptoxis* sequences and a likely misidentified '*Semisulcospira libertina*' sequence (see Xu et al. 2024).

In all phylogenetic trees of this study (*16S*+*COI*, *16S*, *COI*, *COI*-ex, ITS, and mitognomes), the sequences of *S. egretta* sp. nov. clustered in a monophyletic group.

Mitochondrial genome analysis

Seven complete mitochondrial genomes of *Semisulcospira egretta* sp. nov. were sequenced, resulting in circular molecules with lengths ranging from 15,469 bp to 15,677 bp, and have been deposited in NCBI (accession numbers listed in Suppl. material 1: table S5). For clarity in the description, PQ165124 was selected for detailed analysis.

As with other Semisulcospiridae mitogenomes, the nucleotide composition of *S. egretta* sp. nov. exhibits a negative AT-skew (Suppl. material 1: table S9). The newly sequenced mitogenome contains 37 genes (13 PCGs, 2 rRNAs, and 22 tRNAs) (Fig. 5B). Nine of the PCGs and seven tRNAs are encoded on the positive strand (forward strand) and the remaining four PCGs, 15 tRNAs, and two rRNAs are located on the negative strand (reverse strand) (Fig. 5B).

The arrangement of the 13 PCGs and 2 rRNA genes is consistent with all other publicly available cerithioidean mitochondrial genomes (Fig. 5C). The general composition of the GC/AT ratio of the *S. egretta* sp. nov. mitogenome is 37.1/62.9%, PCGs 38.2/61.8%, rRNAs 34.4/65.6%, and tRNAs 36.6/63.4%, with negative GC skew (-0.028) and AT skew (-0.069) (Suppl. material 1: table S9). The AT skew and GC skew of 13 PCGs in *S. egretta* sp. nov. range from -0.126 (*COX2*) to −0.330 (*ATP6*) and 0.073 (*ND2*) to −0. 164 (*ND5*) (Suppl. material 1: table S9).

The total length of the 13 PCGs of *S. egretta* sp. nov. is 11,286 bp and begins with the typical mitogenome codon ATG, except for the ND4 gene which begins with GTG. All stop codons of the 13 PCGs were TAA/TAG. The large ribosomal RNA (16S rRNA) gene of *S. egretta* sp. nov. is 1,344 bp in size, with an A + T content of 66.8%, and its small ribosomal RNA (12S rRNA) is 890 bp, with an A + T content of 63.6% (Suppl. material 1: table S9). The two rRNA genes are located between *trnL1* and *trnV* and *trnT* and *trnS1*. The 22 tRNAs have a size range of 63 bp (*trnC*) to 70 bp (*trnN*) and are 1476 bp in total length (Suppl. material 1: table S9).

The relatively synonymous codon usage (RSCU) values for the PCGs in the mitogenome of the *S. egretta* sp. nov. were analyzed and compared here with *S. gottschei* and *S. coreana*, and showed very high similarity (Supplementary File 7). The most frequently found codons of *S. egretta* sp. nov. include UCU(S), UUA(L), AAA(K), and GUU(V), whereas GCG(A), UCG(S), AAG(K), and ACG(T) have the lowest frequencies (Suppl. material 1: table S9).

Systematics

***Semisulcospira egretta* Y.-Z. Meng, sp. nov.**

https://zoobank.org/A2840C7D-A9ED-48F0-B042-FA0D2DB21F76

Table 2, Figs 1, 6–8

Material examined. ***Holotype.*** China • Haicang District [海沧区], Xiamen City [厦门市], Fujian Province [福建省], China; 24°36'29"N, 117°53'39"E; 13 January 2024; Yuan-Zheng Meng leg.; MBSM-2024-XM-001.

Paratypes. China • 20; the same data as holotype; MBSM-2024-XM-002 to 004, XMU_DYY_TZS_001 to 002 and 005 to 017, YZM/1 (empty shell) • 16; the same locality as the holotype; 16 April 2022; Yuan-Zheng Meng and Li-Wen Lin leg.; XMU_DYY_XMA01 to A15, LWL/1 (empty shell).

Other material. China • 13; Tong'an District [同安区], Xiamen City [厦门市], Fujian Province [福建省], China; 24°52'35"N, 118°03'01"E; 26 February 2023; Yuan-Zheng Meng and Li-Wen Lin leg.; XMU_DYY_XMB01 to B04, XMU_DYY_XMC01 to C06; XMU_DYY_RTM05 and RTM25, YZM/1 (empty shell) • 3; the same locality as holotype; 16 April 2022; Yuan-Zheng Meng and Li-Wen Lin leg.; XMU_DYY_RTM05, XMU_DYY_RTM02, 04 to 05.

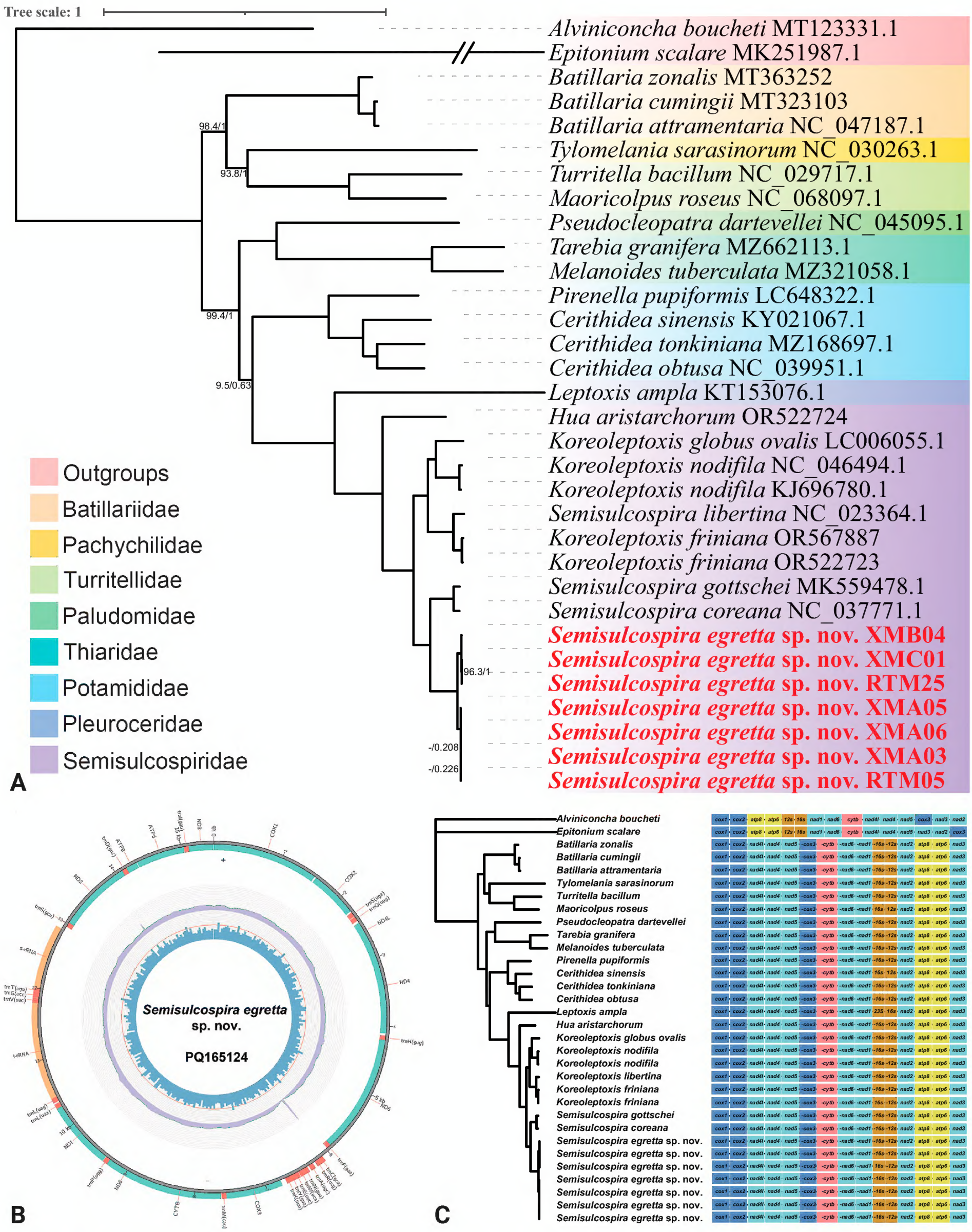

Figure 5. A. Phylogenetic tree of Cerithioidea species inferred from 13 PCGs. The numbers at the internodes represent maximum likelihood (ML) bootstrap (BS) and Bayesian inference (BI) posterior probabilities (PP). The BS equal to 100% and PP equal to 1 were hidden; **B.** Mitochondrial gene map of *Semisulcospira egretta* sp. nov.; **C.** The gene map of the mitochondrion of *S. egretta* sp. nov. XMU_DYY_RTM25 (PQ165125) (**A**) and the gene order of 13 PCGs and 2 rRNAs for Cerithioidea, to show the complete synteny between them (**B**). The GenBank accession numbers used are listed after the species names. Notes: *Semisulcospira egretta* sp. nov. is highlighted in red (**A**).

Table 2. Shell measurements of *Semisulcospira egretta* sp. nov. (n = 45). The holotype is highlighted in red.

Voucher No.	H/mm	B/mm	BW/mm	WA/mm	LA/mm	Whorls	B/H (%)	BW/H (%)	B/BW (%)
MBSM-2024-XM-001	23.0	10.7	15.7	5.2	11.0	6	46.5	68.3	68.2
MBSM-2024-XM-002	16.2	7.6	10.6	3.9	7.5	6	46.9	65.4	71.7
MBSM-2024-XM-003	18.4	9.8	13.1	4.8	9.6	6	53.3	71.2	74.8
MBSM-2024-XM-004	19.4	9.1	12.8	4.5	10.3	6	46.9	66	71.1
MBSM-2024-XM-005	16.2	7.9	11.1	3.8	8.7	6	48.8	68.5	71.2
XMU_DYY_TZS_001	18.8	10.0	12.4	5.1	9.5	7	53.2	66	80.6
XMU_DYY_TZS_002	24.1	11.1	16.6	5.8	12.2	6	46.1	68.9	66.9
XMU_DYY_TZS_005	19.6	9.4	13.8	4.9	10.1	5	48	70.4	68.1
XMU_DYY_TZS_006	17.5	9.5	12.7	4.2	9.2	6	54.3	72.6	74.8
XMU_DYY_TZS_007	15.9	8.7	11.6	4.3	9.6	6	54.7	73	75
XMU_DYY_TZS_008	18.4	8.9	12.7	4.3	9.6	6	48.4	69	70.1
XMU_DYY_TZS_009	21.1	10.2	14.1	4.7	10.5	5	48.3	66.8	72.3
XMU_DYY_TZS_010	20.7	10.1	14.7	4.7	11.0	6	48.8	71	68.7
XMU_DYY_TZS_011	21.4	9.3	14.6	4.7	10.7	6	43.5	68.2	63.7
XMU_DYY_TZS_012	19.9	9.0	13.4	4.4	9.7	5	45.2	67.3	67.2
XMU_DYY_TZS_013	15.6	8.0	10.9	4.1	8.7	6	51.3	69.9	73.4
XMU_DYY_TZS_014	14.0	7.3	10.3	3.5	7.7	5	52.1	73.6	70.9
XMU_DYY_TZS_015	18.7	8.9	13.0	4.4	9.5	6	47.6	69.5	68.5
XMU_DYY_TZS_016	20.1	8.7	13.4	4.3	9.5	6	43.3	66.7	64.9
XMU_DYY_TZS_017	18.2	9.1	12.5	4.3	9.3	6	50	68.7	72.8
XMU_DYY_XMA01	11.9	5.9	8.1	3.5	6.1	6	49.6	68.1	72.8
XMU_DYY_XMA02	11.7	5.7	7.7	3.7	5.9	6	48.7	65.8	74
XMU_DYY_XMA03	11.1	6.0	7.9	3.5	6.0	5	54.1	71.2	75.9
XMU_DYY_XMA04	10.7	6.2	7.1	4.0	5.9	5	57.9	66.4	87.3
XMU_DYY_XMA05	13.1	7.1	9.1	4.3	8.2	6	54.2	69.5	78
XMU_DYY_XMA06	14.0	6.9	9.9	4.4	7.6	6	49.3	70.7	69.7
XMU_DYY_XMA07	13.5	6.5	8.8	3.6	6.5	6	48.1	65.2	73.9
XMU_DYY_XMA08	14.5	7.0	10.0	4.7	7.4	5	48.3	69	70
XMU_DYY_XMA09	15.6	7.6	10.7	4.5	8.2	6	48.7	68.6	71
XMU_DYY_XMA10	15.6	7.6	10.9	4.9	8.3	5	48.7	69.9	69.7
XMU_DYY_XMA11	16.4	8.0	11.3	5.2	8.0	5	48.8	68.9	70.8
XMU_DYY_XMA12	16.6	8.0	11.3	5.2	8.0	5	48.2	68.1	70.8
XMU_DYY_XMA13	17.9	8.0	11.7	4.7	8.3	6	44.7	65.4	68.4
XMU_DYY_XMA14	20.3	9.2	12.8	5.7	9.4	5	45.3	63.1	71.9
XMU_DYY_XMA15	23.1	10.5	15.6	6.2	10.7	5	45.5	67.5	67.3
XMU_DYY_XMB01	11.8	6.0	8.2	3.5	4.8	6	50.8	69.5	73.2
XMU_DYY_XMB02	12.0	6.4	9.1	3.9	7.5	6	53.3	75.8	70.3
XMU_DYY_XMB03	13.4	6.7	9.7	3.4	7.2	6	50	72.4	69.1
XMU_DYY_XMB04	15.8	7.5	11.0	4.4	7.8	7	47.5	69.6	68.2
XMU_DYY_XMC01	14.3	6.2	8.7	3.6	6.6	7	43.4	60.8	71.3
XMU_DYY_XMC02	15.2	6.9	10.0	4.7	7.8	6	45.4	65.8	69
XMU_DYY_XMC03	16.0	6.8	10.2	4.1	6.9	7	42.5	63.8	66.7
XMU_DYY_XMC04	11.3	6.3	7.9	3.2	6.2	6	55.8	69.9	79.7
XMU_DYY_XMC05	14.9	7.2	10.5	3.7	7.6	7	48.3	70.5	68.6
XMU_DYY_XMC06	18.3	7.8	11.2	4.4	8.3	7	42.6	61.2	69.6
Statistical data of the shell measurements (n = 45)									
minimum	10.7	5.7	7.1	3.2	4.8	5.0	42.5	60.8	63.7
maximum	24.1	11.1	16.6	6.2	12.2	7.0	57.9	75.8	87.3
average	16.6	8.0	11.3	4.4	8.4	5.9	48.8	68.4	71.4
standard deviation	3.46	1.46	2.32	0.67	1.65	0.63	3.68	3.05	4.26

Diagnosis. Viviparous. Shell medium size, up to 25 mm, yellow-brownish, elongate, smooth, thin, comprising seven to nine whorls. Pregnant female bears about 150 embryos. Embryonic shell about 0.5 mm in length. Only known in Xiamen, Fujian Province, China.

Description. ***Shell*** (Fig. 6A–F, Table 2, n = 45). Shell height 10.7–24.1 mm, width 5.7–11.7 mm, body whorl height 7.1–16.6 mm, aperture width 3.2–6.2 mm, aperture height 4.8–12.2 mm. Shell elongate, conical, thin, brown to yellowish-brown, sometimes with 1–3 brown bands and black sediment, five to seven whorls. Shell surface smooth, without any spiral ridges or ribs. Apex always eroded. Aperture ovate. Measurements and counts are shown in Table 1.

External morphology (Fig. 1, n = 20). Snout, neck, sides of foot black in color with golden spots on tentacles. Snout broad, squarish, with long cephalic tentacles, tentacles significantly longer than snout in length. Viviparous. Mantle edge smooth.

Operculum (Fig. 6A, n = 17). Corneous, ovate, brown, with approximately 3 whorls. Nucleus of operculum in lower one third of the operculum.

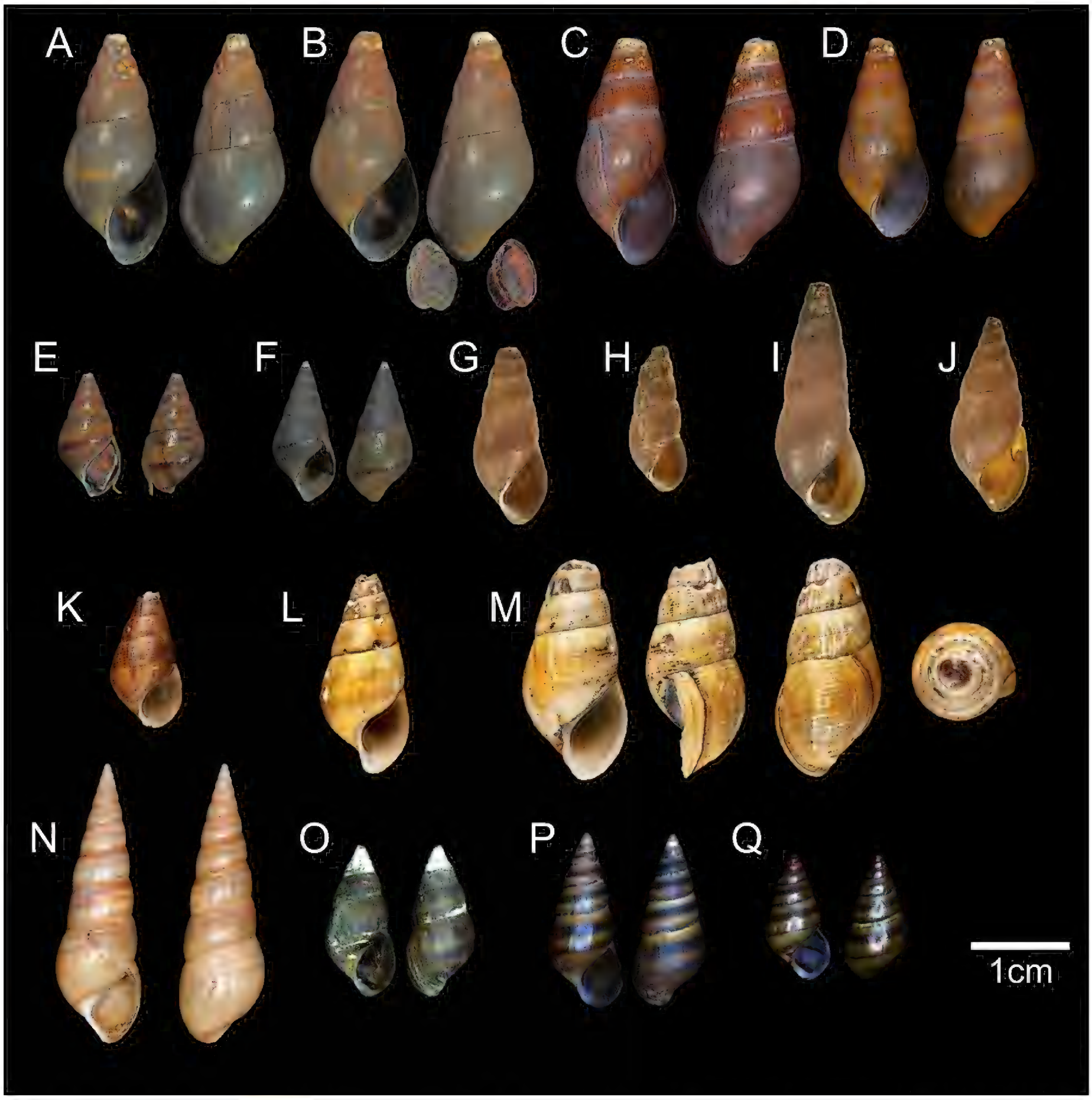

Figure 6. *Semisulcospira egretta* sp. nov. and congeners. A-F, *S. egretta* sp. nov. **A.** Holotype, with operculum, XMU_DYY_TZS_003; **B.** Paratype 6, XMU_DYY_TZS_002; **C.** Paratype 34, XMU_DYY_XMA15; **D.** Paratype 32, XMU_DYY_XMA13; **E.** XMU_DYY_XMB01; **F.** XMU_DYY_XMC01; **G–J.** *S. kurodai*; **G.** ZMB 114713a; **H.** ZMB 114713b; **I.** ZMB 114754a; **J.** ZMB 114754b; **K–M.** *S. libertina*, syntypes of *Melania japonica*, NHMUK ZOO 20210249; **N.** *S. pleuroceroides*, lectotype of *Melania pleurocerides*, MNHN-IM-2000-21289; **O.** *S. guilinensis*, holotype, NNUH20240901; **P.** *S. tigra*, holotype, NNUH20240101; **Q.** *S. sangpuensis*, holotype, NNUH20241001. **A–F.** Photographed by Yuan-Zheng Meng; **G–J.** from Köhler (2016); **O–Q.** from He et al. (2024).

Radula (Fig. 7A, B, n = 9). Central teeth with one large central triangular cusp, and three to four conical denticles on each side, lateral teeth with single prominent triangular cusp, with two to three small denticles on each side, inner marginal teeth with four to five and outer with five to seven flattened, rounded denticles.

Midgut (Fig. 8A, n = 4). Oesophagus opening under ledge on the left side of midgut floor. Marginal fold extending anteriorly from oesophageal aperture alongside major typhlosole, then turning posteriorly bordering right margin of sorting area. Sorting area triangular, posterior tip curving left around sorting area pad. Midgut roof left of sorting area coarsely folded and cuticularised. Gastric shield tongue-shaped and concave. Glandular pad small, triangular, overhanging lip forming shallow pocket behind gastric shield and curving right. Crescentic ridge long, shallow, deep crescentic groove. Caecal fold along midgut floor opposite caecum. Style sac U-shaped, in contact with intestinal groove.

Reproductive systems. Female (Fig. 8B, n = 20). Long narrow oviduct, emerging from ovary, entering near seminal receptacle without protrusion. Ventral edge of spermatophore bursa with groove with curved surface forming sperm gutter, extending toward mantle cavity. Brood pouch uterine, on dorsal side of spermatophore bursa, inflated dorsally, separated into many cells, filled with eggs or embryos, depending on level of sexual development.

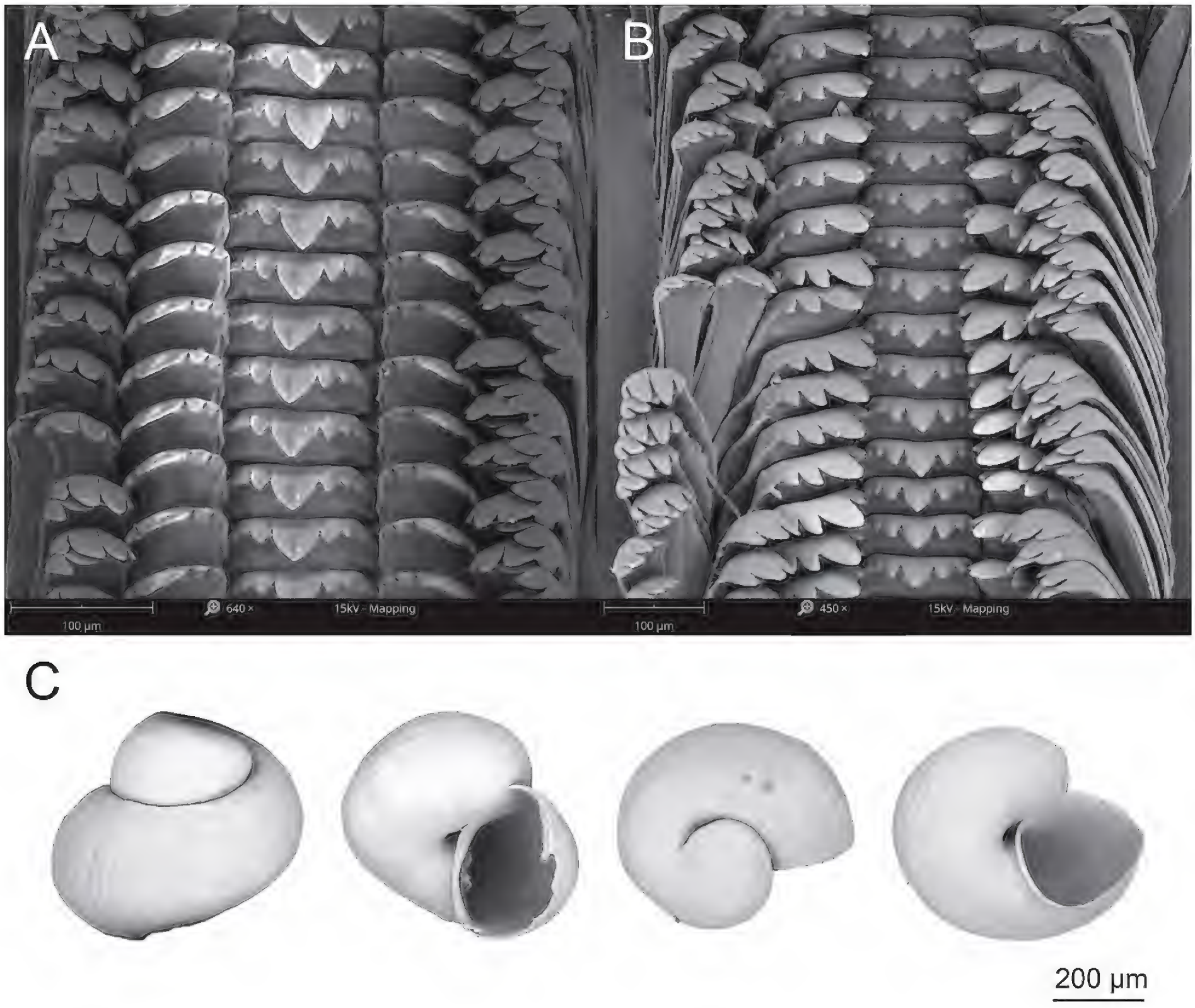

Figure 7. The radulae and embryonic shells of *S. egretta* sp. nov. **A.** The radula image of XMU_DYY_TZS_008; **B.** The radula image of XMU_DYY_TZS_010; **C.** Different views of embryonic shells. (Photographed by Y.-Z. Meng).

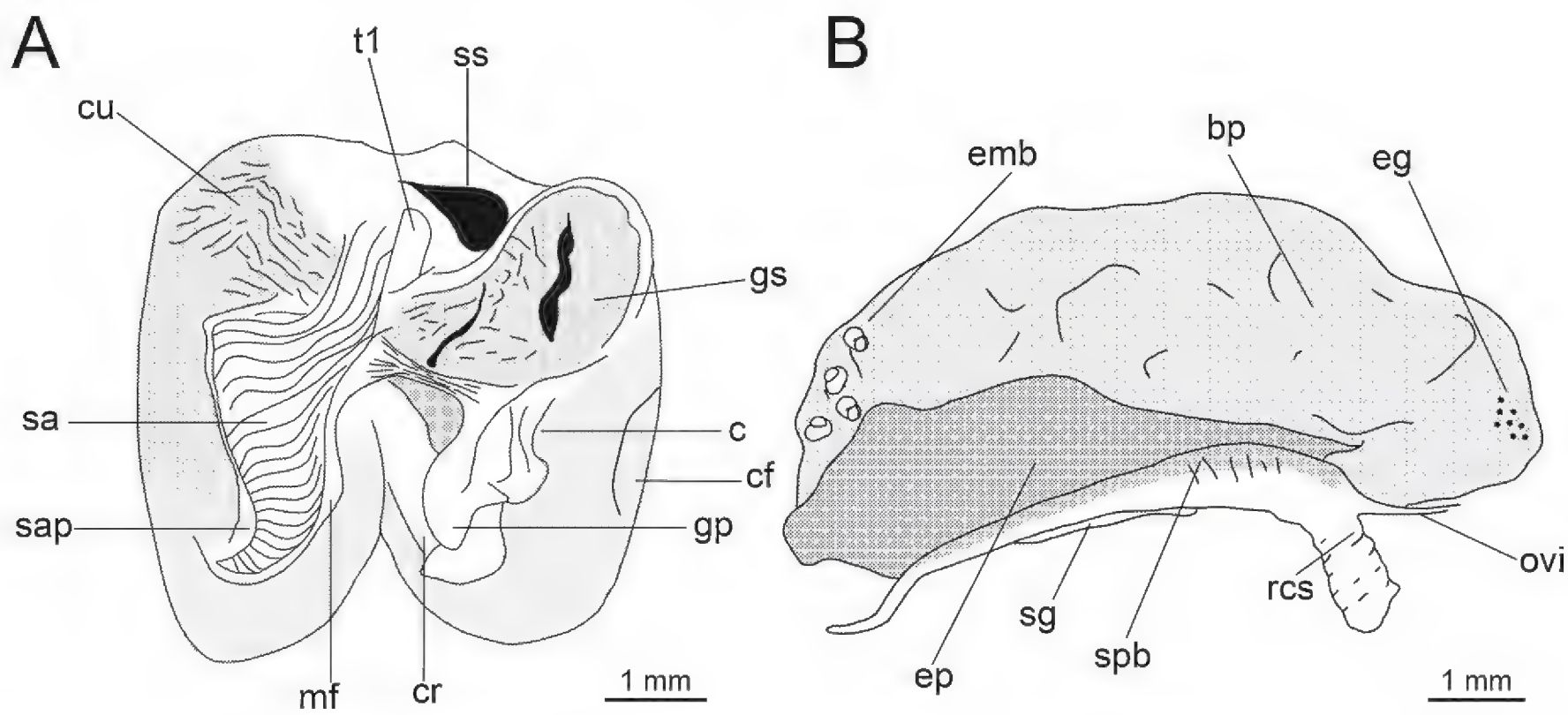

Figure 8. Anatomy of *S. egretta* sp. nov. **A.** Midgut anatomy. **B.** Female reproductive system anatomy. Abbreviations: bp, brood pouch; c, caecum; cf, caecal fold; cr, cresentic ridge; cu, cuticula; eg, egg; emb, embryos; ep, epithelial tissue; gp, glandular pad; gs, gastric shield; mf, marginal fold; ovi, renal oviduct; rcs, seminal receptacle; sa, sorting area; sap, sorting area pad; sg, sperm gutter; spb, spermatophore bursa; ss, style sac aperture; t1, major typhlosole. (Drawn by Yuan-Zheng Meng).

Male. No male was observed.

Sex ratio. All 20 dissected individuals were female.

Embryonic shell (Fig. 7C, D, n = 150). Globose, smooth, possessing slight spiral striae, usually comprising about 1.5 whorls. About 0.5 mm in height.

Habitat and distribution (Fig. 1). The species was collected in mountain streams of Haicang District and Tong'an District of Xiamen, Fujian Province.

Etymology. The word *egretta* means "egret" in Latin. Egret is the city bird of Xiamen, which also suggests the type locality of this species. We suggest the Chinese common name as "白鹭短沟蜷".

Comparative remarks. Many species of *Semisulcospira* have sculptured shells, for example, *S. ningpoensis* (I. Lea, 1857), *S. gredleri* (O. Boettger, 1886), and *S. reticulata* Kajiyama & T. Habe, 1961, etc. But *Semisulcospira egretta* sp. nov. can be easily distinguished from them by its thin and smooth shell. This species is also characterized by the relatively small embryonic shells (about 0.5 mm in length), and the denticle numbers on the inner and outer marginal teeth of its radula (four to five on the inner and five to seven on the outer). For detailed comparisons with similar species (also with relatively smooth shells refer) to Table 3. *Semisulcospira egretta* sp. nov. is the only semisulcospirid species documented in Xiamen; the only region from which it is currently known.

Discussion

The taxonomic status and phylogeny of *Semisulcospira egretta* sp. nov.

Our species hypothesis for *Semisulcospira egretta* sp. nov. is supported by both morphological and molecular data, as follows:

1. *S. egretta* sp. nov. is placed in the genus *Semisulcospira* due to its viviparity, a synapomorphy of the genus (Glaubrecht 2006; Strong and Köhler 2009; Köhler 2017).
2. Some taxonomic characters in *S. egretta* sp. nov., such as the thin, smooth shell (n = 45); embryonic shell size (about 0.5 mm; n = 150); and radular morphology (n = 9) exhibit relative stability and can distinguish it from other congeners. These characters (e.g. reproductive organ, shell, and radula) have been widely adopted by previous studies for classifying semisulcospirid species (Strong and Köhler 2009; Du et al. 2019a, 2019b; He et al. 2024).
3. Intraspecific uncorrected p-distances were 0% to 3.1% (0.84% on average) in *COI*, 0% to 2.2% (0.72% on average) in *16S* and 0% to 1.0% (0.56% on average) in ITS among sequences of *S. egretta* sp. nov. Although the two populations show a relatively large *COI p*-distance with each other (2.7% to 3.1%, 2.8% on average), the distances within the same population are very low respectively (0.077% of location A and 0.12% of location B). In most situations, the *COI p*-distance between congeners of Semisulcospiridae is more than 4% except for *Hua aubryana*, *H. jacqueti* and *H. tchangsii* (Du et al. 2019a). Given the ITS *p*-distance between the two populations is not more than 1%, and there is no significant morphological difference between the two populations, we regard specimens from locations A and B as a single species. The higher distances between the two populations are possibly due to geographical isolation.
4. All phylogenetic trees based on mitochondrial data (*16S*+*COI*, *COI*, *16S*, *COI*-ex, ITS, and mitogenomes) show all *S. egretta* sp. nov. sequences forming a monophyletic group.

Despite evidence from morphology, DNA barcoding, and molecular phylogenetic analysis based on mitochondrial data supporting our proposed new species, many species still lack detailed morphological studies and molecular data, which impedes comprehensive comparisons across the genus, including those reported in China. Furthermore, the putative new species has not been tested using nuclear data. Therefore, further validation of the proposed new species is essential.

In China, our new species may be confused in the literature with *S. libertina*. This species has been documented in many regions of China, including Jilin, Liaoning, Zhejiang, Anhui, Jiangxi, Hubei, Hunan, Fujian, Taiwan, Guangdong, Guizhou, and Yunnan provinces since the last century (Xu 2008). However, these records lack detailed morphological studies and do not record the reproductive mode. Zeng et al. (2015) reported the complete mitogenome of *S. libertina*, but we have suggested it is a misidentification of a *Koreoleptoxis* species (this study, Xu et al. 2024). This makes us question the validity of all *S. libertina* records in China.

We noticed that many individuals marked as the same species (e.g. *S. libertina* or *S. reiniana*) did not form a monophyletic group in the phylogenetic tree (Fig. 3), and appeared in different clades of the tree. Species of *Semisulcospira* did not cluster into a monophyletic group, either. This result has also been observed in some previous studies, such as Lee et al. (2007), Köhler (2016, 2017), and Morita et al. (2024). Some different hypotheses were proposed to explain these phenomena: retention of ancestral polymorphisms, introgression, paternal leakage, heteroplasmy, recombination, pseudogenes, non-neutral mitochondrial evolution, and reversion of reproduction strategies, etc. (Köhler 2016, 2017; Du et al. 2019b). Köhler (2017) suggested that clade A and B should be omitted from the phylogenetic reconstruction of the Semisulcospiridae due to their unusual evolutionary rates. However, it is beyond the focus and indeed the capacity of this study to explain these problems.

Our phylogenetic trees based on short sequences (*16S*, *COI*, *16S*+*COI*, *COI*-ex, and ITS) exhibit low resolution (Fig. 3, Suppl. materials 2–4, 6), making the phylogenetic

Table 3. Summary of differences between *Semisulcospira egretta* sp. nov. and some other congners. Details from He et al. (2024), Davis (1969), Ko et al. (2001), Du et al. (2019a), Du and Yang (2023).

Species	Shell	Whorls	Length of Embryo / mm	Embryo Whorl Numbers	Embryos Number	Inner Marginal Tooth Cusp Number	Outer Marginal Tooth Cusp Number	Distribution
S. calculus	No sculpture, whorls convex.	5-7	~0.4	–	6	6	6	North Korea, and Liaoning Province, China
S. egretta sp. nov.	No sculpture, thin.	5-7	~0.5	1.5	~200	4-5	5-7	Xiamen, Fujian Province, China
S. elonga	No sculpture.	5-7	–	3-4	24-37	6	6	Ningbo, Zhejiang Province, China
S. guilinensis	Growth lines not obvious, thick.	8-9	–	–	–	4	6-7	Guilin, Guangxi Zhuang Autonomous Region, China
S. kurodai	With spiral ribs on apical whorls, fragile.	4-6	< 1.62	< 3.0	18-66	–	–	Japan
S. libertina	With spiral cords, thick.	4-5	< 1.35	< 3.0	137-979	5	6	Japan, Korea, China (doubtful)
S. pleuroceroides	With spiral ribs.	7-14	–	2	117-120	4-6	4-5	The middle and lower reaches of the Yangtze River, China
S. sangpuensis	Surface with some horizontal ribs, thick.	5	–	–	–	7-8	4-5	Shantou, Guangdong Province, China
S. tigra	Growth lines not obvious, thick.	7-8	–	–	–	4	9-10	Liuzhou, Guangxi Zhuang Autonomous Region, China

relationships of *Semisulcospira egretta* sp. nov. unclear. Although the phylogenetic trees based on mitogenomes exhibit higher resolution, they do not indicate the exact phylogenetic position of *S. egretta* sp. nov. due to limited data (Fig. 5A). We suggest phylogenomic methods including transcriptomes, ultraconserved elements (UCEs), whole genomes, and restriction association site DNA (RAD), etc. with broader taxon sampling in studying the phylogeny of Semisulcospiridae, is needed.

Unresolved problems in the phylogenetics of *Semisulcospira*

The phylogenetic trees based on *16S* and *COI* sequences in this study suggested that *Semisulcospira* is polyphyletic (Fig. 3). This result is generally consistent with previous studies (Köhler 2016, 2017; Du et al. 2019a, 2019b). However, these results are only based on a few genes, especially partial mitochondrial genes. It remains to be tested whether *Semisulcospira* is truly polyphyletic when a larger number of nuclear genes and a broader taxon sampling are included. In this study, we do not solve this problem directly but provide transcriptome data for future research.

Authors' contributions

Conceptualization and project administration: D.-Y. Y. and Y.-B. X.; Data curation: Y.-Z. M., S. Zeng and D.-Y. Y.; Formal analysis: S. Zeng, D.-Y. Y and Y.-Z. M.; Funding acquisition: Y.-B. X; Investigation: Y.-Z. M.; Methodology: D.-Y. Y. and Y.-Z. M.; Supervision: D.-Y. Y., C. J. G. and X.-P. Z.; Visualization: S. Zeng; Writing (original draft): Y.-Z. M., D.-Y. Y., Y.-B. X and S. Zeng.; Writing (review & editing): All authors.

Acknowledgments

This study was supported by the Basic Scientific Research Project of Fujian Provincial Public Welfare Research Institutes (2022R1013001). The authors are grateful to Dr. Li-Na Du's advice in dissecting operations and DNA amplification. We are also grateful to Dr. Zhi Wang and Ms. Ying Lu's help with SEM operations. We also thank Mr. Li-Wen Lin, Ying-Lin Shi, and Ms. Cheng Chen's help in collecting specimens, and Mr. Bing-Peng Xing from TIO for accepting our type specimens. We sincerely thank the reviewers and editor for their constructive comments on this article.

References

Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T (2009) trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25: 1972–1973. https://doi.org/10.1093/bioinformatics/btp348

Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, Xia R (2020) TBtools: an integrative toolkit developed for interactive analyses of big biological data. Molecular Plant 13: 1194–1202. https://doi.org/10.1016/j.molp.2020.06.009

Chen H, Shi B-Y, Du L-N, Sun H-Y (2023) Description of a New Species of Hua (Gastropoda: Semisulcospiridae) from Guizhou, China, Based on Morphology and Molecular Evidence. Zoological Science 40: 414–421. https://doi.org/10.2108/zs230025

Chiu Y-W, Bor H, Kuo P-H, Hsu K-C, Tan M-S, Wang W-K, Lin H-D (2017) Origins of *Semisulcospira libertina* (gastropoda: semisulcospiridae) in Taiwan. Mitochondrial DNA Part A 28: 518–525. https://doi.org/10.3109/24701394.2016.1149823

Choi EH, Choi NR, Hwang UW (2021) The mitochondrial genome of an Endangered freshwater snail *Koreoleptoxis nodifila* (Caenogastropoda: Semisulcospiridae) from South Korea. Mitochondrial DNA Part B 6: 1120–1123. https://doi.org/10.1080/23802359.2021.1901626

Davis GM (1969) A taxonomic study of some species of *Semisulcospira* in Japn (Mesogastropoda: Pleuroceridae). Malacologia 7: 211–294.

Du L-N, Yang J-X (2023) Colored Atlas of Chinese Melania. Henan Science and Technology Press, Zhengzhou, 225 pp.

Du L-N, Köhler F, Yu G-H, Chen X-Y, Yang J-X (2019a) Comparative morpho-anatomy and mitochondrial phylogeny of Semisulcospiridae in Yunnan, south-western China, with description of four new species (Gastropoda: Cerithioidea). Invertebrate Systematics 33: 825–848. https://doi.org/10.1071/IS18084

Du LN, Chen J, Yu GH, Yang JX (2019b) Systematic relationships of Chinese freshwater semisulcospirids (Gastropods, Cerithioidea) revealed by mitochondria) sequences. Zoological Research 40: 541–551. https://doi.org/10.24272/j.issn.2095-8137.2019.033

Gim J-A, Baek K-W, Hah Y-S, Choo HJ, Kim J-S, Yoo J-I (2021) Draft genome of *Semisulcospira libertina*, a species of freshwater snail. Genomics & Informatics 19: Glaubrecht M (2006) Independent evolution of reproductive modes in viviparous freshwater Cerithioidea (Gastropoda, Sorbeoconcha)-a brief review. Basteria 70: 23–28. https://doi.org/10.5808/gi.21039

Glaubrecht M (2006) Independent evolution of reproductive modes in viviparous freshwater Cerithioidea (Gastropoda, Sorbeoconcha)-a brief review. Basteria 70: 23–28.

Guo Y, Fu Z, Feng J, Ye Y, Li J, Guo B, Lv Z (2019) The complete mitochondrial genome and phylogenetic analysis of *Epitonium scalare* (Gastropoda, Epitoniidae). Mitochondrial DNA Part B 4: 1070–1071. https://doi.org/10.1080/23802359.2019.1586462

Hartnell College Genomics Group, Andrade P, Arreola L, Belnas M, Bland E, Castillo A, Cisneros O, Contreras V, Diaz C, Do KT (2019) The complete mitogenome of the invasive Japanese mud snail *Batillaria attramentaria* (Gastropoda: Batillariidae) from Elkhorn Slough, California, USA. Mitochondrial DNA Part B 4: 4031–4032. https://doi.org/10.1080/23802359.2019.1688719

He YM, Lu YZ, Fu ZY, Xiang HQ, Chen H (2024) Description of 17 new species of Semisulcospiridae (Gastropoda: Cerithioidea) from southern China based on morphological and molecular evidence. Ecologica Montenegrina 75: 12–32. https://doi.org/10.37828/em.2024.75.2

Hilgers L, Grau JH, Pfaender J, von Rintelen T (2016) The complete mitochondrial genome of the viviparous freshwater snail *Tylomelania sarasinorum* (Caenogastropoda: Cerithioidea). Mitochondrial DNA Part B 1: 330–331. https://doi.org/10.1080/23802359.2016.1172046

Jin J-J, Yu W-B, Yang J-B, Song Y, DePamphilis CW, Yi T-S, Li D-Z (2020) GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. Genome Biology 21: 1–31. https://doi.org/10.1186/s13059-020-02154-5

Jombart T, Kendall M, Almagro-Garcia J, Colijn C (2017) treespace: Statistical exploration of landscapes of phylogenetic trees. Molecular Ecology Resources 17: 1385–1392. https://doi.org/10.1111/1755-0998.12676

Kalyaanamoorthy S, Minh BQ, Wong TK, Von Haeseler A, Jermiin LS (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. Nature Methods 14: 587–589. https://doi.org/10.1038/nmeth.4285

Kato S, Itoh H, Fukumori H, Nakajima N, Kanaya G, Kojima S (2022) The mitochondrial genome of the threatened tideland snail *Pirenella pupiformis* (Mollusca: Caenogastropoda: Potamididae) determined by shotgun sequencing. Mitochondrial DNA Part B 7: 632–634. https://doi.org/10.1080/23802359.2022.2060143

Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular Biology and Evolution 30: 772–780. https://doi.org/10.1093/molbev/mst010

Kim YK, Lee SM (2018) The complete mitochondrial genome of freshwater snail, *Semisulcospira coreana* (Pleuroceridae: Semisulcospiridae). Mitochondrial DNA Part B 3: 259–260. https://doi.org/10.1080/23802359.2018.1443030

Ko J, Lee J, Kwon O (2001) Study on radulae of seven species of the Family Pleuroceridae in Korea. The Korean Journal of Malacology 17: 105–115.

Köhler F (2016) Rampant taxonomic incongruence in a mitochondrial phylogeny of *Semisulcospira* freshwater snails from Japan (Cerithioidea: Semisulcospiridae). Journal of Molluscan Studies 82: 268–281. https://doi.org/10.1093/mollus/eyv057

Köhler F (2017) Against the odds of unusual mtDNA inheritance, introgressive hybridisation and phenotypic plasticity: systematic revision of Korean freshwater gastropods (Semisulcospiridae, Cerithioidea). Invertebrate Systematics 31: 249–268. https://doi.org/10.1071/IS16077

Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. Molecular Biology and Evolution 35: 1547. https://doi.org/10.1093/molbev/msy096

Lee SY, Lee HJ, Kim YK (2019) Comparative analysis of complete mitochondrial genomes with Cerithioidea and molecular phylogeny of the freshwater snail, *Semisulcospira gottschei* (Caenogastropoda, Cerithioidea). International Journal of Biological Macromolecules 135: 1193–1201. https://doi.org/10.1016/j.ijbiomac.2019.06.036

Lee W-K, Hou BK, Ju S-J, Kim S-J (2020) Complete mitochondrial genome of the hydrothermal vent provannid snail *Alviniconcha boucheti* (Gastropoda: Abyssochrysoidea) from the North Fiji Basin. Mitochondrial DNA Part B 5: 1848–1849. https://doi.org/10.1080/23802359.2020.1750983

Lee T, Hong HC, Kim JJ, Foighil DÓ (2007) Phylogenetic and taxonomic incongruence involving nuclear and mitochondrial markers in Korean populations of the freshwater snail genus *Semisulcospira* (Cerithioidea: Pleuroceridae). Molecular Phylogenetics and Evolution 43: 386–397. https://doi.org/10.1016/j.ympev.2007.02.018

Letunic I, Bork P (2024) Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. Nucleic Acids Research 52(W1): W78–W82. https://doi.org/10.1093/nar/gkae268

Ling Y, Zhang Y, Ngatia JN, Zhou H (2022) The complete mitochondrial genome of *Melanoides tuberculata* (Müller, 1774) in Guangdong, China. Mitochondrial DNA Part B 7: 1319–1320. https://doi.org/10.1080/23802359.2022.2054735

Liu Y, Zhang W, Wang Y, Wang E (1979) Economic Fauna of China: Freshwater Mollusca. Science Press, Beijing, 134 pp.

Liu Y-Y, Zhang W-Z, Wang Y-X, Duan Y-H (1994) Eight New Species of Freshwater Molluscs in Southwest China. Acta Zootaxonomica Sinica 19: 25–36.

Lydeard C, Cummings KS (2019) Freshwater Mollusks of the World: a Distribution Atlas. JHU Press, Baltimore, 242 pp. https://doi.org/10.1353/book.66164

Miura O, Urabe M, Mori H, Chiba S (2020) Ancient drainage networks mediated a large-scale genetic introgression in the East Asian freshwater snails. Ecology and Evolution 10: 8186–8196. https://doi.org/10.1002/ece3.6523

MolluscaBase eds (2024) MolluscaBase. https://www.molluscabase.org [accessed September 3.2024]

Morita K, Saito T, Uechi T, Sawada N, Miura O (2024) Out of the ancient lake: Multiple riverine colonizations and diversification of the freshwater snails in the genus *Semisulcospira* around Lake Biwa. Molecular Phylogenetics and Evolution 191: 107987. https://doi.org/10.1016/j.ympev.2023.107987

Nguyen L-T, Schmidt HA, Von Haeseler A, Minh BQ (2015) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Molecular Biology and Evolution 32: 268–274. https://doi.org/10.1093/molbev/msu300

Nguyen DH, Lemieux C, Turmel M, Nguyen VD, Mouget J-L, Witkowski A, Tremblay R, Gastineau R (2018) Complete mitogenome of *Cerithidea obtusa*, the red chut-chut snail from the Cần Giờ Mangrove in Vietnam. Mitochondrial DNA Part B 3: 1267–1269. https://doi.org/10.1080/23802359.2018.1532832

Paradis E, Schliep K (2019) ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics 35: 526–528. https://doi.org/10.1093/bioinformatics/bty633

R Core Team (2023) R: A language and environment for statistical computing. In: R Foundation for Statistical Computing (Ed.) Vienna, Austria.

Ronquist F, Teslenko M, Van Der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Systematic Biology 61: 539–542. https://doi.org/10.1093/sysbio/sys029

Sawada N, Nakano T (2021) Revisiting a 135-year-old taxonomic account of the freshwater snail *Semisulcospira multigranosa*: designating its lectotype and describing a new species of the genus (Mollusca: Gastropoda: Semisulcospiridae). Zoological Studies 60: e7.

Sawada N, Fuke Y, Miura O, Toyohara H, Nakano T (2024) Redescription of *Semisulcospira reticulata* (Mollusca, Semisulcospiridae) with description of a new species from Lake Biwa, Japan. Evolutionary Systematics 8: 127–144. https://doi.org/10.3897/evolsyst.8.124491

Steenwyk JL, Buida III TJ, Gonçalves C, Goltz DC, Morales G, Mead ME, LaBella AL, Chavez CM, Schmitz JE, Hadjifrangiskou M (2022) BioKIT: a versatile toolkit for processing and analyzing diverse types of sequence data. Genetics 221: iyac079. https://doi.org/10.1093/genetics/iyac079

Stelbrink B, Kehlmaier C, Wilke T, Albrecht C (2019) The near-complete mitogenome of the critically endangered *Pseudocleopatra dartevellei* (Caenogastropoda: Paludomidae) from the Congo River assembled from historical museum material. Mitochondrial DNA Part B 4: 3229–3231. https://doi.org/10.1080/23802359.2019.1669081

Strong EE (2011) More than a gut feeling: utility of midgut anatomy in phylogeny of the Cerithioidea (Mollusca: Caenogastropoda). Zoological Journal of the Linnean Society 162: 585–630. https://doi.org/10.1111/j.1096-3642.2010.00687.x

Strong EE, Köhler F (2009) Morphological and molecular analysis of '*Melania*' *jacqueti* Dautzenberg & Fischer, 1906: from anonymous orphan to critical basal offshoot of the Semisulcospiridae (Gastropoda: Cerithioidea). Zoologica Scripta 38: 483–502. https://doi.org/10.1111/j.1463-6409.2008.00385.x

Strong EE, Garner JT, Johnson PD, Whelan NV (2022) A systematic revision of the genus *Juga* from fresh waters of the Pacific Northwest, USA (Cerithioidea, Semisulcospiridae). European Journal of Taxonomy 848: 1–97. https://doi.org/10.5852/ejt.2022.848.1993

Whelan NV, Strong EE (2016) Morphology, molecules and taxonomy: extreme incongruence in pleurocerids (Gastropoda, Cerithioidea, Pleuroceridae). Zoologica Scripta 45: 62–87. https://doi.org/10.1111/zsc.12139

Xu X (2008) Classification and Morphology on Pleuroderidae (Gastropoda: Mesogastropoda) of Some Areas in Middle Reaches of Changjiang River. MA. Sc Thesis, Nanchang, Jiangxi Province, China: Nanchang University, 79 pp.

Xu Y, Luo P, Wang P, Zhu P, Zhang H, Wu H, Liao Y, Yu M, Fu J (2019) The complete mitochondrial genome of *Cerithidea sinensis* (Philippi, 1848). Mitochondrial DNA Part B 4: 2742–2743. https://doi.org/10.1080/23802359.2019.1644549

Xu Y, Zeng S, Meng Y, Yang D, Yang S (2024) The mitochondrial genome of *Hua aristarchorum* (Heude, 1889)(Gastropoda, Cerithioidea, Semisulcospiridae) and its phylogenetic implications. ZooKeys 1192: 237–255. https://doi.org/10.3897/zookeys.1192.116269

Yan C, Feng J, Ye Y, Li J, Guo B (2020a) The complete mitochondrial genome and phylogenetic analysis of *Batillaria cumingi* (Gastropoda: Batillariidae). Mitochondrial DNA Part B 5: 2355–2356. https://doi.org/10.1080/23802359.2020.1772685

Yan C, Feng J, Ye Y, Li J, Guo B (2020b) The complete mitochondrial genome and phylogenetic analysis of *Batillaria zonalis* (Gastropoda: Batillariidae). Mitochondrial DNA Part B 5: 2256–2257. https://doi.org/10.1080/23802359.2020.1772139

Yang S, Deng Z (2022) The complete mitochondrial genome of *Cerithidea tonkiniana* (Mabille, 1887) in Guangxi, China. Mitochondrial DNA Part B 7: 669–670. https://doi.org/10.1080/23802359.2021.2006813

Yang D, Zeng S, Wang Z, Zhang Y, Yang D, Glasby CJ, Hwang JS, Cai L (2024) Molecular systematics of *Perinereis* and an investigation of the status and relationships of the cultured species *Perinereis wilsoni* Glasby & Hsieh, 2006 (Annelida, Nereididae). Zoosystematics and Evolution 100: 1297–1314. https://doi.org/10.3897/zse.100.127201

Yin N, Zhao S, Huang X-C, Ouyang S, Wu X-P (2022) Complete mitochondrial genome of the freshwater snail *Tarebia granifera* (Lamarck, 1816)(Gastropoda: Cerithioidea: Thiaridae). Mitochondrial DNA Part B 7: 259–261. https://doi.org/10.1080/23802359.2022.2026832

Zeng T, Yin W, Xia R, Fu C, Jin B (2015) Complete mitochondrial genome of a freshwater snail, *Semisulcospira libertina* (Cerithioidea: Semisulcospiridae). Mitochondrial DNA 26: 897–898. https://doi.org/10.3109/19401736.2013.861449

Zeng L, Wang Y, Zhang J, Wu C (2016) Complete mitochondrial genome of *Turritella terebra bacillum*. Mitochondrial DNA Part B 1: 350–351. https://doi.org/10.1080/23802359.2016.1144088

Zhang D, Gao F, Jakovlić I, Zou H, Zhang J, Li WX, Wang GT (2020) PhyloSuite: An integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. Molecular Ecology Resources 20: 348–355. https://doi.org/10.1111/1755-0998.13096

Supplementary material 1

Additional tables

Authors: Yi-Bin Xu, Yuan-Zheng Meng, Sheng Zeng, Hang-Jun Wang, Shen Zhong, De-Yuan Yang, Xi-Ping Zhou, Christopher J. Glasby

Data type: xlsx

Explanation note: **table S1.** Specimen information of *Semisulcospira egretta* sp. nov. **table S2.** Protocols for extracting DNA sequences. **table S3.** Single gene dataset used for phylogenetic analysis. **table S4.** The COI gene dataset used for phylogenetic analysis in fig. S1. **table S5.** Mitochondrial genome dataset used for phylogenetic analysis. **table S6.** Original and trimAl lengths of the PCGs and PCGsAA sequences. **table S7.** Best partitioning schemes and models based on different datasets for Bayesian inference (BI) and maximum likelihood (ML) analysis. **table S8.** Pairwise distances and Kimura 2 Parameters between sequences of Clade D in Fig. 3. **table S9.** Mitochondrial genome analysis for mitogenomes of Semisulcospiridae.

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/zse.101.136882.suppl1

Supplementary material 2

Phylogenetic tree inferred from *16S* and *COI* sequences, based on different trimming methods

Authors: Yi-Bin Xu, Yuan-Zheng Meng, Sheng Zeng, Hang-Jun Wang, Shen Zhong, De-Yuan Yang, Xi-Ping Zhou, Christopher J. Glasby

Data type: pdf

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/zse.101.136882.suppl2

Supplementary material 3

Phylogenetic tree inferred from *16S* sequences, based on different trimming methods

Authors: Yi-Bin Xu, Yuan-Zheng Meng, Sheng Zeng, Hang-Jun Wang, Shen Zhong, De-Yuan Yang, Xi-Ping Zhou, Christopher J. Glasby

Data type: pdf

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/zse.101.136882.suppl3

Supplementary material 4

Phylogenetic tree inferred from *COI* sequences, based on different trimming methods

Authors: Yi-Bin Xu, Yuan-Zheng Meng, Sheng Zeng, Hang-Jun Wang, Shen Zhong, De-Yuan Yang, Xi-Ping Zhou, Christopher J. Glasby

Data type: pdf

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/zse.101.136882.suppl4

Supplementary material 5

The topological differences of *16S*+*COI*, *16S*, and *COI* datasets generated by TreeSpace

Authors: Yi-Bin Xu, Yuan-Zheng Meng, Sheng Zeng, Hang-Jun Wang, Shen Zhong, De-Yuan Yang, Xi-Ping Zhou, Christopher J. Glasby

Data type: pdf

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/zse.101.136882.suppl5

Supplementary material 6

Maximum Likelihood tree inferred from extensive *COI* sequences (*COI*-ex)

Authors: Yi-Bin Xu, Yuan-Zheng Meng, Sheng Zeng, Hang-Jun Wang, Shen Zhong, De-Yuan Yang, Xi-Ping Zhou, Christopher J. Glasby

Data type: pdf

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/zse.101.136882.suppl6

Supplementary material 7

Codon usage analysis in *S. egretta* sp. nov. and congeners

Authors: Yi-Bin Xu, Yuan-Zheng Meng, Sheng Zeng, Hang-Jun Wang, Shen Zhong, De-Yuan Yang, Xi-Ping Zhou, Christopher J. Glasby

Data type: pdf

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/zse.101.136882.suppl7